DSIC Introduction Video Script

Slide 1:
Hi. I am Julie Wietzke, the Information and Outreach Manager at the caBIG® Data Sharing and Intellectual Capital Knowledge Center. I am going to provide you with an introduction to the Data Sharing and Intellectual Capital (or DSIC) Workspace and Knowledge Center as well as the Data Sharing & Security Framework or DSSF.

Slide 2:
The DSIC Workspace is a strategic level caBIG® Workspace, which means that it is focused on overarching issues and designed to develop policies and guidelines that support other workspaces and the broader caBIG® community. As a strategic Workspace, it is also involved in the development and refinement of the caBIG® strategic plan. The mission of the DSIC Workspace is to facilitate data sharing between and among caBIG® participants by addressing legal, regulatory, policy, proprietary, and contractual barriers to data exchange. The primary goals of the Workspace are to develop recommendations for policies, procedures, and best practices, prepare white papers and comment letters on proposed policies and guidelines, develop problem scenarios that illustrate issues confronted by caBIG® participants, support reviews of caBIG® tools under development, and provide education and outreach to caBIG® participants and relevant offices at their institutions like IRBs and technology transfer offices. The members of the DSIC Workspace include biomedical researchers, clinicians, technology transfer experts, intellectual property and regulatory attorneys, policy specialists, patient advocates, bioethicists and bioinformaticists. The Workspace works directly with the DSIC Knowledge Center to help set priorities.

Slide 3:
The Enterprise Support Network (ESN) is a collection of diverse organizations that, with guidance from NCICB, caBIG® Program staff, and caBIG® workspaces, expand and support the caBIG® community by providing appropriate services, mentoring and expertise. Critical functions of the Enterprise Support Network include:

* Enabling further progress toward achieving the caBIG® program goal of establishing and extending interoperability,
* Supporting the needs of individuals and institutions of the current caBIG® community, and broader cancer and biomedical research communities that wish to adopt caBIG® standards, applications, and infrastructure,
* Servicing a range of audiences, including IT administrators, end users, and senior decision makers, and
* Augmenting the existing capacity of the caBIG® Workspaces

There are 2 critical components of the ESN: Support Service Providers, which are independent, NCI-approved organizations that provide client-specific caBIG® support under negotiated client-provider business arrangements, and Knowledge Centers, which I will explain in more detail on the next slide.

Slide 4:
Knowledge Centers have been established at institutions with demonstrated expertise in a specific area of focus or domain of interest to caBIG® . There are currently six Knowledge Centers. Services supported by these centers include:

* Web-based support for education, outreach, training, and deployment needs related to their domain to the caBIG® and broader cancer enterprise,
* Community outreach to expand interest in caBIG® tools and compatibility,
* Domain expertise for those interested in integrating caBIG® into their programs, and
* Maintenance of a central repository of domain-specific tools, documentation, policies, and standards

The 6 Knowledge Centers are:
* caGrid
* Clinical Trials Management Systems:
* Molecular Analysis Tools
* Tissue/Biospecimen Banking and Technology Tools
* Vocabulary

and our Knowledge Center, the Data Sharing and Intellectual Capital Knowledge Center.

Slide 5:
The caBIG® Data Sharing and Intellectual Capital (DSIC) Knowledge Center is an NCI-supported entity led by the University of Michigan. The DSIC Knowledge Center provides a centralized, authoritative repository of processes, model agreements, and other resources to encourage and facilitate data sharing to advance scientific discovery, consistent with applicable legal, regulatory, ethical and contractual requirements.
Key services provided by the Knowledge Center include:
* Domain expertise for other caBIG® Knowledge Centers and Workspaces,
* Analysis and summary of relevant laws, regulations, policies and standards and their likely impact on data sharing initiatives,
* Development of decision support and analytic tools,
* Authorship, publication and dissemination of white papers, and
* Coordination with the caBIG® community and external groups on data sharing initiatives

Slide 6:
The Knowledge Center makes use of a wiki as a repository for our tools and resources including the Data Sharing and Security Framework or DSSF, the associated decision support tools, a model informed consent and related

materials, as well as various knowledge article and bibliographies. More detail will be provided later on the tools and resources.

We also maintain forums, which are designed to be the venue for members of the caBIG® community and other interested parties to provide feedback to the Knowledge Center on our tools and resources or to suggest additional tools that would be helpful to the community. The Forums are also the place for individuals to pose questions to the Knowledge Center or post items of potential interest to others in the community. Anyone who registers and creates an account can participate in the Forums. We encourage you to register if you haven't already.

Slide 7:
Here is a screen shot of the DSIC KC Wiki.

Slide 8:
Here is a screen shot of the Forums.

Slide 9:
The Knowledge Center is co-directed by Elaine Brock, Sr. Associate Director of the Division of Research Development and Administration at the University of Michigan, and Rachel Nosowsky, Senior Counsel at Miller, Canfield Paddock & Stone. Alex Kanous is the Operations Manager. As I stated in the beginning, I am the Outreach and Information Manager. Kevin Smith is the Informatics Liaison for the Knowledge Center.

Slide 10:
Here is contact information for the Knowledge Center including our 800# and the urls for our wiki and forums. We encourage you to post any comments or questions to the Forums or feel free to call us on the 800# if you prefer.

Slide 11:
We will now move into the DSSF or Data Sharing & Security Framework, but first I am going to give you some background.

Slide 12: no audio

Slide 13:
Biomedicine in general -- and cancer research especially -- are moving in a clear direction towards an enhanced understanding of the molecular basis of disease. We are witnessing the transition to personalized medicine where we will soon see unique *molecular level* characteristics of the individual patient and disease driving the prevention of disease and the delivery of health care. This paradigm will require the synthesis of multidimensional data and the joining of multiple, diverse communities, including researchers, care providers and data repositories, in order to direct care that is based on the molecular characteristics of disease.

•Making progress with the broad range of new technologies now available requires an integrated approach, which you may have heard described as translational research or personalized medicine. In order to realize the benefits of the molecular medicine revolution, including earlier detection, more productive drug discovery and development, more individualized patient care, we need to utilize all the new technologies available in basic and clinical research.

•Biomedical informatics is a critical enabler and integrator, facilitating the seamless flow of data from bench to bedside . . . The associated infrastructure, tools and applications allow researchers to break down silos and foster collaboration within and across the research-care continuum, thus enhancing the ability to deliver personalized, targeted, evidence-based medicine.

Slide 14:
caBIG® arises from this movement as a network that enables the cancer community to share data and knowledge and thereby accelerate the discovery of new approaches for the detection, diagnosis, treatment, and prevention of cancer, ultimately improving patient outcomes.

It is important to note that the caBIG® vision is not constrained to cancer, but instead extends to all biomedical research and patient-centric approaches to molecular medicine.

Slide 15:
Enabling the data sharing objectives of caBIG® is caGrid, or “The Grid,” an underlying network infrastructure that provides the basis for connectivity between and among cancer community institutions and caBIG® tools.

Slide 16:
The Grid operates by allowing users to query a Grid Index service to discover the metadata about relevant stored data that has been advertised by host institutions.

Slide 17:
Here is a diagram of the interaction of the various components of the caGrid infrastructure.

Slide 18:
Now let’s turn to the Data Sharing and Security Framework.

The Framework has its genesis from the fact that since the inception of the caBIG® pilot, NCI recognized that technical solutions alone were insufficient: The DSIC WS was chartered at the beginning of the caBIG® pilot project to address potential or perceived barriers to developing and using the caBIG® infrastructure to share data. Its membership is by necessity diverse.

The DSSF emerged as DSIC's flagship product from the caBIG® pilot: a response to the often paralyzing effect on researchers and their institutions of addressing multiple sources of potential restrictions on data sharing; the Framework provides a consistent and analytic approach to navigating these issues and enabling the flow of biomedical research data.

Slide 19:

Widespread data sharing is only possible if the issues that arise from it are overcome. The natures of these roadblocks take the form of technological, legal, and social issues.

Technical constraints in sharing data are resolved through the suite of tools that have been developed by the caBIG® community.

Slide 20:
Legal issues are addressed through the tools and policies developed within the DSIC WS and KC, including the DSSF.

Slide 21:
Social issues restricting the dissemination of data are tackled by educating institutions on the value of data sharing and providing incentives for doing so.

Slide 22:
Only by addressing each of these areas of constraint can we achieve the goal of data accessibility.

Slide 23:
We recognize that many of the challenges of cross-institutional data sharing are not technical. They include:

•Varying obligations under federal and state privacy and security laws
•Oversight of human subjects research by ethical review boards or IRBs where local requirements vary substantially based on interpretations of applicable requirements
•Academic considerations such as the need to secure grants and publish results in peer reviewed literature
•Researcher, institutional and sponsor concerns re: protection of intellectual property
•Patient safety concerns related to premature access to un-validated information as well as public perceptions regarding privacy, security and confidentiality of electronic health data

Slide 24:
To mitigate the challenges to data sharing, caBIG®® uses:
- a federated architecture that enables local control of research and clinical data;

- an analytical framework to encourage continuous, consistent analysis of legal, regulatory, ethical and proprietary barriers to data sharing and identify solutions; and
- standards, tools and infrastructure – broadly available -- to facilitate appropriate data sharing to support biomedical research and personalized medicine while preventing prohibited or unethical access

It as been recognized from the beginning of the caBIG® Pilot: data sharing and security are tightly coupled.
• the development of the caBIG® infrastructure is predicated on the fundamental concept of federation.
• caBIG® software and resources are widely distributed, interlinked, and available to everyone in the cancer research community, but institutions can maintain local control over their own resources and data using appropriately secured mechanisms.

Slide 25:
Next, I would like to give an overview of the Framework. We are going to focus on two aspects:

First, the vertical columns, or strands, represent four areas of sensitivity in four domains.

Then, within each area, the Framework helps identify three levels of sensitivity: low, medium and high, coded, respectively, green, yellow and orange.

The purpose of the Framework is to provide the decision logic for analyzing the sensitivity of data to determine the appropriate type of data sharing mechanism to employ and the appropriate levels of security and data access controls to apply to access the data.

•It is based on the ***sensitivity*** of data and access controls appropriate to the levels of sensitivity of the data.

•It recognizes that there are varying levels of sensitivity of health information and that many data exchanges require agreements, validation of users, authorization of intended uses and so forth; institutions determine who is authorized and under what conditions (levels of security, types of contractual restrictions, etc.) In particular, **some data are highly sensitive and should never be shared without individual permission.**

•The Framework is designed to eliminate or reduce actual and perceived barriers to data sharing, i.e., moving data from the orange to the yellow to the green lane, thus spurring scientific progress and eventually the cure for cancer, [and ultimately other diseases]. However, because the caBIG® infrastructure is premised on the concept of federation, individual entities that control access to

data are responsible for assessing the risk and consequent protection required for any data set to be shared. The Data Sharing and Security Framework permits the user (researcher and/or institutional professional) to consistently analyze any constraints associated with sharing data and specimens.

Slide 26:
The organization arrives at an overall level of sensitivity for the data by weighting the outcomes of the four elements according to its own judgment.

The outcome, that is, a low, medium or high sensitivity rating, determines how the organization wants to control access to that data.

Slide 27:
The organization offering to share data determines the controls on access to that data by determining:
•The level of certainty needed regarding the authentication of the identity of data users, and
•Whether particular authenticated groups or individuals are authorized to access the particular data.

The levels of security attached to data sensitivities of various levels are informed by guidance from the National Institute of Standards and Technology (NIST).

Slide 28:
The Framework helps Providers select the type of data sharing mechanism that best fits their needs based on how they assess the sensitivity of the data to be shared. The Data Sharing and Security Framework is not a strict policy or guideline, but instead a means of analysis.

Slide 29:
Here is a decision tree-like arrangement of the Framework, which is just a different way of looking at the framework.
•The purple diamonds represent the areas of sensitivity that must be analyzed to come to an understanding of the nature of any mandatory legal restrictions on data sharing and agreements, if any, necessary to facilitate proposed exchanges
•Answering the questions posed in the purple diamonds will lead you to a determination of the appropriate type of data sharing mechanisms to employ and the data access controls to apply

This concludes this particular presentation. Thank you for your time. I hope this presentation has been informative. Please see our wiki for presentations on other topic areas.